

# METHOD AND APPARATUS FOR INITIATING DATA CAPTURE IN A DIGITAL CAMERA BY TEXT RECOGNITION

5

## FIELD OF THE INVENTION

The present invention relates generally to digital cameras and, more specifically, to a method and associated apparatus for initiating data capture in a digital camera by the recognition of text patterns such as words or phrases.

10

## BACKGROUND OF THE INVENTION

Digital cameras provide powerful features and convenience to hobbyists and professionals alike. In applications such as genealogy, for example, genealogists often search through large books or stacks of records for relevant information. Once relevant information is found on a particular page, a genealogist may take a photograph of the page using a digital camera.

Digital photography provides several advantages over traditional silver-emulsion-film photography. First, a user can see immediately on the display of the digital camera whether or not the photograph is acceptable. Secondly, the photograph is created directly in digital format and can be used immediately in printed reports, photo albums, Web pages, or e-mail messages. Although the foregoing advantages are significant, the task of searching through the pages of a book or a stack of records looking for relevant information can still be tedious and time consuming. Often, the search centers on finding a specific word such as a surname or an important phrase. Unfortunately, prior-art digital cameras do not aid the user in finding text patterns on a printed page. The task of searching for relevant pages is left entirely to the user.

It is thus apparent that there is a need in the art for an improved method or apparatus for locating relevant pages of printed material and digitally photographing them easily and efficiently.

5

## SUMMARY OF THE INVENTION

According to one aspect of the invention, a method is provided for initiating data capture in a digital camera upon recognition of a text pattern. In one embodiment, the method comprises a manual two-pass approach. In another embodiment, the method is highly automated. According to another aspect of the invention, a digital camera for carrying out the method is also provided.

10

Other aspects and advantages of the present invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating by way of example the principles of the invention.

15

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a functional block diagram of a digital camera in accordance with the present invention.

FIG. 2 is an illustration of a text-input user interface device for a digital camera in accordance with one embodiment of the present invention.

20

FIG. 3 is a flowchart of the operation of the digital camera shown in FIG. 1 in accordance with one embodiment of the present invention.

FIG. 4 is a flowchart of the operation of the digital camera of FIG. 1 in accordance with another embodiment of the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

FIG. 1 is a functional block diagram of a digital camera 100 in accordance with one embodiment of the present invention. In FIG. 1, user interface 105 comprises input device 110, display 115, and optional audible tone generator 120.

5 Input device 110 sends commands to controller 125 over data bus 130 to specify the modes of operation of digital camera 100 and acts as a source for text input. User interface 105 provides feedback to users via display 115 and optional audible tone generator 120. Imaging device 135 converts an optical image received from optical system 140 to a digital image. The digital image may be stored temporarily in  
10 optional random access memory (RAM) 145. Optical character recognition (OCR) module 150 converts digital images of printed material to text. Image compression module 155 compresses digital images to reduce their storage requirements. Non-volatile memory 160 provides long-term storage of digital representations. Throughout this detailed description, the term "digital representation" will be used  
15 to denote a text file, a digital image (compressed or uncompressed), a message digest, check sum, or any other digital representation of a target scene or any combination thereof.

In a typical implementation, imaging device 135 comprises a charge-coupled device (CCD), an analog-to-digital converter (A/D), a gain control, and a digital  
20 signal processor (DSP), as is well known in the art (not shown in FIG. 1). Input device 110 typically comprises one or more buttons for selecting modes or acquiring text input. For example, FIG. 2 shows one method in which a text pattern may be acquired in digital camera 100, which may not include an alphanumeric keypad. Hereinafter, the term "text pattern" will be used to denote a textual string

comprising one or more alphanumeric characters, including spaces. Scroll buttons 205 and 210 control which portion of alphanumeric character list 215 is visible on display 220. Cursor 225 indicates which alphanumeric character is currently selectable. Pressing select button 230 adds the currently highlighted character to text pattern 235 to be searched. The same process of scrolling using scroll buttons 205 and 210 and selecting characters using select button 230 is repeated for each character in text pattern 235 until the entire text pattern has been specified.

Optionally, user interface 105 may include a microphone and a speech recognizer for accepting voice input (not shown in FIG. 1). One example of how voice input may be incorporated in a digital camera may be found in published European Patent Application EP-1014338A1. Yet another method for acquiring a text pattern in digital camera 100 is to retrieve a previously saved text pattern. For example, a text pattern may be read from a text file, or the previous text pattern for which a search was performed may be loaded by default until overridden by new input. Audible tone generator 120 is an optional component of digital camera 100 and has been included in FIG. 1 for the purpose of illustration only. In some implementations, a flashing or persistent light, a vibrator, or other method of providing feedback to users may be preferable to an audible tone. Non-volatile memory 160 is typically flash memory. In some implementations, non-volatile memory 160 may be of the removable type, such as a memory stick or magnetic disk. Optical character recognition module 150 may be implemented in hardware, software, firmware, or a combination thereof. Image compression module 155, which may be implemented in hardware, software, firmware, or a combination thereof, typically produces images in the popular JPEG format.

FIG. 3 is a flowchart illustrating the operation of digital camera 100 in accordance with one exemplary embodiment of the present invention. At 305, controller 125 determines whether or not text-recognition mode has been activated. If so, a text pattern to be recognized is acquired at 310 by one or more of the methods previously described. A target scene at which to aim digital camera 100 is selected at 315. In typical applications, the target scene is a page in a book or other printed material compatible with OCR technology. Upon receipt of a continue command from user interface 105 at 320, imaging device 135 produces a digital image, which may be stored temporarily in RAM 145, at 325. OCR module 150 converts the digital image to a text file at 330. The text file may also be stored temporarily in RAM 145. At 335, controller 125 searches for the text pattern that was acquired at 310. If the text pattern is recognized, an audible tone is emitted at 340. Upon receipt of a confirmation command at 345, imaging device 135 produces a refined digital image at 350 of whatever target scene is currently being received from optical system 140. The purpose of this arrangement is to provide both a quick search to determine whether a text pattern is present on a printed page and a subsequent follow-up second pass in which a user may align digital camera 100 more precisely to take a final photograph of the page. However, the confirmation command at 345 may be replaced by different techniques in other embodiments of the present invention. For example, the refined digital image may be captured automatically after a predetermined delay following the recognition of the text pattern at 335.

Controller 125 determines at 355 in what format output data should be saved in non-volatile memory 160. Depending on which format or formats have been

specified, image compression (e.g., for JPEG output), OCR, or other format conversion may need to be performed on the refined digital image at 355. At 360, a digital representation of the target scene is saved in non-volatile memory 160. At 365, controller 125 checks for a termination command. If none is received, control  
5 returns to 315, where a new target scene is selected. If a termination command is received, digital camera 100 exits text-recognition mode at 370. If the text pattern is not recognized at 335, control proceeds to 365.

The embodiment described in connection with FIG. 3 exemplifies a manual two-pass approach to the process of text recognition and picture taking.

10 Specifically, the continue command at 320 and the confirmation command at 345 are intended to give the user maximum control over the final digital representation saved, but such a manual approach is by no means the only manner in which the present invention may be embodied. For example, a greater degree of automation may be introduced to speed up the taking of multiple photographs and to provide  
15 greater convenience to the user.

FIG. 4 is a flowchart illustrating the operation of digital camera 100 in accordance with another embodiment of the present invention. The method of operation shown in FIG. 4 incorporates a higher degree of automation than that shown in FIG. 3. In FIG. 4, text-recognition mode is activated at 305, a text  
20 pattern to be recognized is acquired at 310, and a target scene is selected at 315, as described in connection with FIG. 3. As soon as text-recognition mode becomes active, or, optionally, after a brief predetermined delay, imaging device 135 produces at 325 a temporary digital image of whatever target scene is currently being received from optical system 140. Note that controller 125 does not wait for

a continue command in this particular embodiment. The temporary digital image is converted to a text file at 335, and controller 125 searches the text file for the text pattern that was acquired at 310. If the text pattern is not recognized, control proceeds to 365. If no terminate command is received at 365, control returns to  
5 315. Otherwise, text-recognition mode is exited at 370.

If the text pattern is recognized at 335, a test is performed at 410 to ensure that the current target scene is different from the previous target scene for which a digital representation was saved in non-volatile memory 160. Such an approach allows the operations at 315, 325, 330, 335, and 410 to be performed in a loop,  
10 thereby freeing the user from the need to indicate via user interface 105 when to take each successive photograph. Thus, in the embodiment associated with FIG. 4, digital camera 100 automatically determines when a new target scene is present and responds by capturing a digital image and saving a digital representation of the target scene in non-volatile memory 160, when the text pattern acquired at 310 is  
15 recognized. Hereinafter, a target scene for which a digital representation has been saved will be referred to as a "saved target scene." There are many ways apparent to those skilled in the art in which the test at 410 may be implemented. In the particular embodiment associated with FIG. 4, the text files associated with the current and the previous saved target scenes are compared. Examples of  
20 comparison techniques include comparing the position of the text pattern acquired at 310 in the current text file to its position in the text associated with the previous saved target scene. Another solution is to compare the text immediately surrounding the text pattern acquired at 310 in the two text files. Other solutions include identifying and comparing page numbers within the text files or comparing

the first few words of the text in the two text files. Still another solution is to compare message digests of the two text files computed using, for example, a digital hashing function such as MD5. In some implementations, direct comparison of the digital images associated with the current and the previously saved target scenes  
5 may be advantageous.

If the text pattern at 410 is recognized for the first time since the activation of text-recognition mode, there is no previous text with which to compare, so control automatically proceeds to 340, where an audible tone is emitting signaling that the text pattern has been recognized. From 355, control proceeds as explained  
10 in connection with FIG. 3.

The present invention may be modified to combine elements of manual confirmation and automation in ways different from that indicated in FIG. 3 and FIG. 4 without departing from the principles of the invention.

The foregoing description of the present invention has been presented for the  
15 purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and other modifications and variations may be possible in light of the above teachings. The embodiments were chosen and described in order to best explain the principles of the invention and its practical application to thereby enable others skilled in the art to best utilize the invention in  
20 various embodiments and various modifications as are suited to the particular use contemplated. It is intended that the appended claims be construed to include other alternative embodiments of the invention except insofar as limited by the prior art.